(19) Japan Patent Agency

(12) **Public Patent Report**

(11) Patent Application Public No. 8-336387
(43) Date open to public: December 24, 1996

(21) Application No. 7-145005
(22) Application Date: June 12[th,] 1995

(54) **[Invention Title]** Sugar chain-elongating protein and DNA derived from yeast *Pichia* genus

(57) **[Summary]**
[Composition] Yeast strain *Pichia*-derived DNA and its encoding protein that is involved in elongating sugar chain to glycoprotein. Production method of glycoprotein characterized by using modified *Pichia* yeast strain in which part of nucleotide sequence is modified so that it decreases the activity of elongating sugar chains.
[Usefulness] By this invention, it is possible by using *Pichia* strain as a host to provide the method of suppressing at the genetic level the intrinsic glycochain-elongating activity in order to produce glycoproteins with the sugar chain of physiologically identical or similar structure. It is possible by using this *Pichia* strain to purify the glycoprotein with glycosyl structure that is identical or similar to mammalian counterparts in ER core glycosylation.

[Content of Patent Request]

[Application item 1] *Pichia*-derived protein involved in sugar chain elongation of glycoprotein that possesses the following amino acid sequence in N-terminal region.
[Formula I]

[Application item 2] *Pichia*-derived protein involved in sugar chain elongation of glycoprotein (as is described in Application item 1) that has the following full-length amino acid sequence.
[Formula II]

[Application item 3] DNA with nucleotide sequence that encodes glycochain-elongation protein described in application item 1 and 2.

[Application item 4] DNA with the following nucleotide sequence that encodes glycochain-elongation protein described in application item 3.
[Formula III]

[Application item 5] Partially-modified DNA (described in application item 3 and 4) that results in the reduction of translating its gene product.

[Application item 6] DNA described in application item 5 in which transformation marker was inserted.

[Application item 7] DNA with the transformation markers (described in item 6) that corresponds to any of the followings: SUC2 gene derived from yeast, HIS4 gene derived from yeast *Pichia*, ARG4 gene, URA3 gene, and G418 antibiotic gene.

[Application item 8] Modified yeast Pichia strains that possess either of DNAs described in item 5 – 7.

[Application item 9] Production method of glycoproteins characterized by use of modified yeast Pichia strains as the host cells.

[Application item 10] Glycoproteins produced by the method (described in the item 9) that includes soluble high-affinity IgE Receptor alpha chain (sFcεRIα), chymase, prourokinase-annexinV fution protein, urea trypsin inhibitor, and IGF1-binding protein 3 (IGF1BP3).

[Detailed explanation of the invention]
[0001]
[Industrial area of application] This invention corresponds to the protein and its gene involved in the glycosylation of proteins that is derived from yeast Pichia strain. Pichia can be widely used as a host of useful expression system to produce genetically-engineered proteins. This protein mediates the elongation of sugar chain, preferably the addition of α-1, 6 mannose chain. This invention is also concerned with its modified DNA, modified Pichia strains that possess it, and the production methods of glycoproteins by use of these Pichia strains as the hosts. Most of the clinically important physiological proteins are glycoproteins, and, by use of these methods, it is possible to purify the glycoproteins with only Man_8 $GlcNAc_2$, an ER core sugar chain that is identical between yeast and mammal. Therefore, thesemodified Pichia strains are useful for producing pharmacologically important glycoproteins.

[0002]
[Problem the current technology and invention are attempting to solve] Most of the proteins functioning in the living body are glycoproteins, which are modified by sugar chains. Up to now the structures of many proteins have been solved by the new analytical methods using HPLC, NMR, AND FAB-MAS as well as the conventional analysis using lectins. On the other hand, a number of researchers vigorously study the functional analysis of sugar chains and found that they play important roles in cell-cell recognition, molecular identification, structural maintenance of proteins, contribution to their activities, clearance, secretion, and localization in the living organisms.

[0003] For example, it is reported that in well-studied human erythropoietin its asparagine-binding sugar chain plays important roles (see Figure 1); its tip part Neu5Ac (sialic acid) participates in clearance in blood, its branched part Gal/GalNAc group (Galactose/N-acetylgalactosamine) in structural contribution to receptor binding, and its core region in the activity of the peptide and various functions [Reference: Takeuchi *Tanpakushitsu Kakusan Kouso* (1992) special issue "Complex glucides" 37, 1713]. Thus sugar chains are not only structurally complicated but functionally versatile, and it is critical to study their structure and functions especially when new glycoproteins are developed as medicines.

[0004] Production of materials by using microorganisms have several advantages over that by using mammalian cells in terms of its cost and the development of culture technique. However, the problem is that microorganisms cannot append the same sugar chains as human counterparts (ex. sugar chains of erythropoietin described above). While mammalian cell-derived glycoproteins include complex groups as shown in Figure 1, asparagine-binding groups with mixed and three high-mannose types, and a variety of mucin type sugar chains, there is no addition of sugar chains in prokaryotes such as E. coli, and there is only limited glycosylation in eukaryotic microorganism, yeast Saccharomyces cerevisiae; asparagine-binding sugar chains

consists of only high-mannose type and the mucin type contains just mannose as the major component. Therefore microorganisms are not useful to produce genetically-engineered glycoprotein production whose sugar chains play an important role, and this explains why Chinese hamster ovary (CHO) cells are currently used to produce erythropoietin.

[0005] To the contrary, research has started to be carry out so as to add structurally identical or similar sugar chains as in mammalian cells by using genetically-engineered yeast strains. In 1994, Schmientek et al. reported success in expressing human-derived β-1,4-galactosyltransferase gene in yeast [Schwientek T and Ernst JR *Gene* 145 299 (1994)]. Krezdrn et al. also study the same way and expressed the activity of β-1,4-galactosyltransferase and α-2,6-sialyltransferase [Krezdrn CH et al. *Eur J Biochem* 220 809 (1994)].

[0006] On the other side, attempts have been made to change the sugar chains themselves on yeast-derived glycoproteins. Yeast high-mannose type chains are hyper-mannosylated --- it contains more mannoses than mammalian counterparts, and has characteristic α-1,6-mannose binding (that elongate at the α-1,3-mannose group that binds to β-mannose) and α-1,3-mannose binding (that is added to the external sugar chain) [See Figure 2].

[0007] In 1992, Jigami's group succeeded in cloning of OCH1 gene (expressing α-1,6-mannosyltransferase) which has been thought to be a key enzyme in elongation of α-1,6-mannose [Nakayama K *EMBO J* 11, 2511 (1992), see Figure 2]. In the mutant strain that lacks OCH1 gene (Δoch1), glycoproteins possess three types of glycochains --- Man_8 ClcNAc$_2$, Man_9 GlcNAc$_2$, Man_{10}GlcNAc$_2$; Man_8 ClcNAc$_2$ is identical to the structure of the ER core glycosyl chain that is shared in mammalian cells (Indicated "Ma" in Figure 2), and Man_9 GlcNAc$_2$ / Man_{10}GlcNAc$_2$; Man_8 ClcNAc$_2$ are those with α-1,3-mannose addition to this ER core structure [Nakanishi-Shindo Y... Jigami Y (1994) *J Biol Chem*]. Furthermore, by making Δoch1/Δmnn1 double mutant to block the transfer of α-1,3-mannose at its tip, it was possible to produce only Man_8 ClcNAc$_2$ that is structurally the same as the ER core sugar chain seen in mammalian cells. Thus the Δoch1/Δmnn1 double mutant is thought to be a valuable host to produce mammalian-derived glycoproteins with high-mannose type sugar chains.

[0008] Recently, methylotrophic yeast, Pichia genus (ex. Pichia pastoris) has been paid attention as a useful host to express heterologous proteins. Pichia yeast is widely used because its culture technique is relatively established and its expression level of secreted proteins is higher than other yeast strains. Little research has been done, however, to understand the mechanism of glycosylation of the protein and the structure of sugar chains in the Pichia strain.

[0009] The purpose of this invention is to provide the Pichia-derived protein and gene that is involved in sugar chain elongation. This is the first report of Pichia-derived protein or gene with such function to be provided. This invention also supplies the production method of glycoprotein by using modified Pichia yeast strains as hosts where the current DNA is engineered so that its expression level is reduced or its ability of glycosylation is blocked. These Pichia strains will be a useful tool to produce glycoproteins with the same or similar sugar chain structure as in mammal by genetic engineering technique.

[0010]
[Means to solve the problems] We inventors work on producing various physiological proteins by using Pichia yest as a host. One problem is most of these proteins are glycoproteins and

glycosylation by yeast is different from that by mammalian cells. By taking various approaches to solve this problem, we succeeded in cloning of the gene that encodes sugar chain elongation protein in the Pichia strain. This invention was completed when we confirmed that this protein is involved in the addition of sugar chain to glycoprotein in the expression system of Pichia yeast.

[0011] This invention is concerned with this protein and its encoding DNA. It is also concerned with the modified DNA in which the expression of its gene product is repressed and the modified DNA in which a transformation marker is inserted. In addition, it is concerned with the modified Pichia yeast strains in which activity of sugar chain elongation is suppressed and the production method in which these modified strains are used as hosts.

[0012] Below is the detailed explanation of the current invention
(1) Protein that is involved in the addition of sugar chains to glycoproteins
This protein is originally expressed by yeast Pichia strain, involved in the first step of glycosylation, and controls the whole glycosylation step.

[0013] These Pichia strains refers to but are not limited to the followings:
Pichia pastoris, Pichia finlandica, Pichia trehalophila, Pichia koclamae, Pichia membranaefaciens, Pichia opuntiae, Pichia thermotolerans, Pichia salictaria, Pichia guercuum, Pichia pijperi
We preferably used Pichia pastoris (referred to as P. pastoris below).

[0014] This protein is derived from yeast Pichia strains, and, though not limited to this if it has the same function above, has the N-terminal amino acid sequence as indicated in Formula I and full-length amino acid sequence as in Formula II.

[0015]
 [Formula I]

[0016]
 [Formula II]

[0017] The patent pending is expected to be effective even if part of the sequence is modified (ex. replacement, deletion, insertion or addition of peptide) as long as it keeps the function described above.

[0018] Since its primary amino acid sequence indicated in Formula II has a high identity (~40%) with α-1,6-mannosyltransferase derived from S. cerevisiae, and its encoding OCH1 gene has a high identity (~55%) with that from S. cerevisiae, it is most likely the Pichia homolog of a key enzyme in α-1,6-mannose addition.

[0019] The protein can be produced by the conventional method where Pichia yeast is cultured at appropriate temperature and the protein is extracted and purified from the cultured yeast. The protein can also be produced by synthesizing polypeptides or by commonly-used genetic engineering techniques, based on the nucleotide and amino acid sequence reported here.

[0020]
(2) DNA encoding the glycochain addition protein
The reported DNA includes nucleotide sequence that encodes the Pichia-derived glycosylation protein as described above. The DNA petitioned refers to any possible nucleotide sequence that

can encode the glycosylation protein described here. The sample sequence is indicated in Formula III below.

[0021]
[Formula III]

[0022] The said DNA can be produced by the common methodology. For example, it is possible to synthesize part or all of DNA by using DNA synthesizer, or to amplify the DNA by the PCR method as chromosome DNA from Pichia yeast (ex. P. pastoris) as a template.

[0023] The said DNA is the first reported Pichia-derived gene encoding the protein that is involved in glycosylation of glycoproteins. It is extremely useful to elucidate the mechanism and structure of sugar chains of glycoproteins in the Pichia yeast strain.

[0024] The protein functions as an additional transfer of a-1,6-mannose to the core sugar chain in Pichia strain, which results in hypermannosylation compared with that in mammalian cells. The genetic elucidation of glycosylation protein provides the better method of expressing glycoproteins with the same or similar sugar chains as those found in mammalian cells. This was achieved by modifying the DNA sequence so that the expression of its gene product is suppressed.

[0025]
(3) Glycosylation DNA that is modified from "Natural" glycosylation DNA
The current invention refers to the Pichia-derived DNA that is partially modified so that the expression of its gene product is repressed.

[0026] "Its gene product" contains any protein that shares the same function as described above. The function refers to the activity of synthesizing and adding sugar chains to glycoproteins, more specifically, of transferring α-1,6-mannose to the core sugar chain. Also, "repression" means not only the absence of expression of the current gene, but the reduced expression of the concerned gene (which results in reduced activity of glycosylation).

[0027] Thus the DNA petitioned for the patent applies to any DNA that results in the absence of glycosylation activity or its decreased activity. For example, it may indicate the DNA where at least one nucleotide is deleted, inserted, or replaced in the "natural glycosylation DNA." This causes the frame shift or amino acid change, resulting in the absence of or altered function from that of the "natural" gene product.

[0028] One of the appropriate modifying method is to insert the transformation marker within the glycosylation DNA encoding region. This makes it much easier to perform the screening of positive mutants that possess the desired DNA by using the marker as a probe. It is also possible to insert the gene that encodes a glycoprotein which you hope to produce. Thus the DNA modification of glycosylation gene and expression of desired glycoprotein to be produced can be achieved at the same time.

[0029] Transformation markers, for instance, include P. pastoris or S. cerevisiae-derived HIS4 gene, ARG4 gene, URA3 gene, SUC2 gene, and G418 antibiotic gene. Among these, we prefer HIS4 gene. Examples of glycoproteins to be produced include but are not limited to: soluble IgE receptor α chain (sGcϵRIα, Patent No. 6-169776), interferon α (Patent No. 61-185189),

urokinase (Patent No. 60-180591), chymase [Caughey GH et al. *J Biol Chem* 266, 12956 (1991)], urea trypsin inhibitor [Kaumeyer JF et al. *Nucleic Acids Res* 14, 7839 (1986)], IGF-binding protein (IGF1BP3, Patent No. 3-505397).

[0030]
(4) Modified yeast Pichia strains
The concerned modified Pichia strain is a yeast strain in which activity of sugar chain elongation is blocked. These strains possess modified gene of glycosylation so that its wild-type gene expression is reduced or completely blocked.

[0031] These modified Pichia strains can be prepared by various methods. For example, selection of mutants in which glycosylation activity is blocked can be made by modifying the original gene or by causing spontaneous mutation in wild-type Pichia. The former method is achieved by introducing the transforming DNA into the specific region in the original gene of glycosylation. The endogenous DNA is replaced with this transformed DNA. Transformed DNA include deletion of the targeted gene, and insertion of selective marker gene or ectopic gene-expressing DNA fragment. This transformation is designed to cause homologous recombination and affect the activity of glycosylation.

[0032] The method of transforming the wild-type Pichia and its culture methods are those widely accepted in these areas. For example, transformation technique include "Spheroplast method" [Creggh et al., *Mol Cell Biol.* 5, 3376 (1985), US Patent No. 4,879,231] and "Lithium chloride method" [Ito et al., *Agric. Biol. Chem.*, 48, 341 (1984), European Patent Application No.312,934, US Patent No. 4,929,535].

[0033] Pichia-derived host cells for the transformation include (but not not limited to) methylotrophic yeast which can utilize methanol as a sole energy source. To name a few, they are auxotropic P. pastoris GTS115 (NRRL Y-15851), P. pastoris GS190 (NRRL Y-18014), P. pastoris PPF1 (NRRL Y-18017), and wild-type P pastoris (NRRL Y-11430, NRRL Y-11431).

[0034] In addition, the preferable strains are those which lack auxotrophic gene; for example, HIS4-deleted P. pastoris GS115 (ATCC20864), ARG4-deleted P. pastoris GS190, HIS4/URA3-deleted P. pastoris GS4-2, HIS4/URA4-deleted P. pastoris PPF1 (NRRL Y-18017). In this cases, it is plausible to choose as a marker DNA an auxotrophic gene that is missing in these strains. This can assure the easy and quick selcetion of the transformed Pichia strains.

[0035] These modified Pichia strains maintain the same growth characteristic as the wild-type in the "natural" culture medium [for example, YPD medium (1% yeast extract, 2% peptone, 2% glucose), YPM medium (1% yeast extract, 2% peptone, 2% methanol)]. This indicates that the modification of glycosylation gene does not affect the growth of Pichia yeast, establishing these modified Pichia yeast as an excellent host for production of pharmacologically useful glycoproteins.

[0036] The claimed materials produced by these modified strains refer to any proteins with sugar chain that have identical or similar structure as in human, but preferably those that are pharmacologically and physiologically important ones. These include soluble high-affinity IgE receptor alpha chain (sFCεRIα), epidermal growth factor (EGF), growth hormone-releasing factor (GRF), IGF1-binding protein 3 (IGF1BP3), pro-urokinase, annexin V fusion protein, chymase, and urea-type trypsin inhibitor.

[0037] Expression system for glycoprotein production can be constructed by various methods. For example, you can introduce glycoprotein-encoding DNA in modified Pichia strain described above, insert the transformation marker along with the glycoprotein-encoding DNA, mutate the glycosylation gene, or transform both the sugar cahing addition gene and glycoprotein-encoding gene at the same time in the wild-type Pichia.

[0038] These Pichia strains for expression system possess at least (1) promotor region, (2) DNA encoding the desired glycoprotein and (3) transcriptional terminator region directionally in the reading frame. These DNA are aligned so that glycoprotein DNA is functionally expressed into RNA.

[0039] The promotor include: P.pastoris-derived AOX1 promotor (for primary alcohol oxidase gene), P.pastoris-derived AOX2 promotor (for secondary alcohol oxidase gene), P.pastoris-derived DAS promotor (for dihydroxyacetone synthase gene), P.pastoris-derived P40 promotor (for P40 gene), P.pastoris-derived promotor for aldehyde dehydrogenase, and P.pastoris-derived promotor for folic acid dehydrogenase gene. Our favorite ones are AOX1 promotor (Ellis et al., *Mol Cell Biol*, 5, 111 (1985), US Patent 4,955,231 etc.) and, more prefarably, mutated AOX2 promotor (Ohi, H et al., *Mol Gen Genet* 243, 489 (1994), Patent No. 4-299984) that was engineered to increase the expression efficiency.

[0040] It is also a good idea to possess the secretion signal sequence in front of the glycoprotein-encoding gene. This technique makes the isolation and purification of the desired glycoproteins easier. The secretion signal sequence include, for example, alpha mating factor (αMF) leader sequence-encoding DNA (including processing region-encoding sequence, Lys-Arg) and fetal lysozyme C signal sequence functioning in methylotrophic yeast.

[0041] The transcriptional terminator refer to any subsegment that signal the transcriptional termination of the DNA that was started by the promotor. It can be a part of the gene encoding the glycoproteins.

[0042] The current expression system may contain the selective marker gene in addtion to the DNA sequence described above. These include HIS4, ARG4, URA3, S. cerevisiae SUC2, and G418 antibiotic gene.

[0043] The modified Pichia strains that are transformed into the desired phenotype can produce glycoprotein by culturing them in the commonly way. The common methods include (but are not limited to) culturing in YPD and YPM medium. The culturing temperature is 20 – 30 °C, preferably, 23 – 25 °C, which is the optimum temperature for growing Pichia yeast. Optimum pH of the medium is selected based on the preference of the yeast strain. Furthermore ventilation and/or shaking is applied when necessary. After the culture, the supernatant is obtained and the desired protein is purified from the supernatant by the public methods; these include, for instance, fractionation technique, and ion-exchange, gel-filtration, hydrophobic, affinity-column chromatography.

[0044] This invention provides, for the first time, Pichia-derived protein and gene that is involved in glycosylation of proteins. This provides the useful tool to understand the mechanism of sugar chain binding and addition to the glycoprotein. This also presents the methodology of engineering Pichia yeast at the genetic level in order to express and produce glycoproteins with

identical or similar character as those pharmacologically and physiologically significant. These Pichia strains have the same growth capability as the parental strain, and reduced or absent activity of sugar chain elongation compared with the parental strain. Thus they can produce glycoproteins with the ER core chain structure that is shared by both yeast and mammalian cells, making them powerful hosts for producing glycoproteins.

[0045]
[Working Example] We will explain the current invention in detail by giving practical examples. This invention, however, is not limited to these described below. The plasmid, restriction enzyme, T4 DNA ligase, and other materials are available commercially and used in the conventional way. The techniques including DNA cloning, determination of nucleotide sequence, transformation of host cells, culture of transformed cells, extraction of enzyme from cultured medium, and purification are those that are widely acknowledged in the area or available in the literature.

[0046] Working example 1: Obtaining the glycosylation gene derived from Pichia yeast
We searched for the Pichia-derived DNA involved in glycosylation by amplifying Saccharomyces cerevisiae-derived glycosylation gene, *OCH1* by PCR method and by performing Southern analysis of genes from Pichia strains with *OCH1* as a probe.

[0047] (1) Acquisition and amplification of S. cerevisiae-derived OCH1 gene by PCR method
In order to clone OCH1 gene that is involved in glycosylation in S. cerevisiae, we made use of the DNA sequence available in publication [The EMBO Journal vol.11 no.7 p2511-2519 (1992): p2513, Fig.2]. Based on the sequence, we chemically synthesized primers which are complementary to the either end of the open reading frame and attached with HindIII recognition site, N-terminal primer (5'-CGAAGCTTATGTCTAGGAAGTTGTTCCCACCTG-3') and C-terminal primer (5'-CGAAGCTTATTTATGACCTGCATTTTTATCAG-3'), by using the DNA synthesizer (ABI Inc., model 392 DNA/RNA synthesizer). By using these primers and DNA prepared from S. cerevisiae AH22 strain (based on the method in: Sherman, F., Fink, G.R. and Hicks, J.B. (1986) Laboratory course manual for methods in yeast genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), we performed PCR reaction (94°C 1min, 50°C 2 min, 72°C 2min / 25 cycle) [DNA Thermal Cycler Model PJ2000, Perkin-elmer Inc.]. The amplified DNA fragment was separated on agarose gel electorophoresis, and we confirmed the single band on the gel. It showed an expected size of 1458 bp.

[[0048] (2) Subcloning of S. cerevisiae-derived *OCH1* gene
After digesting the PCR-amplified fragment obtained in (1) with HindIII, it was subcloned into HindIII site in pUC19 vector. The constructed plasmid (pKM049, See Fig. 3) was digested with several restriction enzymes (BamHI, EcoRI, Kpn I) and found to correspond to the restriction map of OCH1 gene [EMBO J. 11, 7 p2511-2519 (1992): p2512, Fig. 1 and p2513, Fig. 2].

[0049] (3) Southern hybridization of Pichia-derived chromosomal DNA with OCH1 gene as a probe
The Pichia strain (Pichia pastoris GTS115 strain) was grown at 30°C for 3 days in YPD medium (1% yeast extract, 2% peptone, 2% glucose). Chromosomal DNA was prepared using the method by Sherman et al. [Sherman F, Fink GR, and Hicks JB (1986) Laboratory course manual for methods in yeast genetics, Cold Spring Harbor Laboratory, New York]. The resultant DNA was restriction digested with various enzymes, electorophoresed on the agarose gel, and transferred to the nylon membrane (Hybond-N, Amersham Inc.). We prepared a probe from HindIII-digested OCH1 gene described in (1) by using "DIG-ELISA labeling kit" (Boehringer-Mannheim Inc.).

Using this probe, we performed Southern hybridization by the conventional method [Shambrook J, Fritsh EF, and Maniatis T (1989) Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory, New York], and investigated if there is any gene in the Pichia strain that is homologous to *OCH1*.

[0050] Since the homology between S. cerevisiae-derived OCH1 gene and its Pichia counterpart is unkonwn, we examined various conditions including hybridization temperature (65˚C, 55˚C, 45˚C) and washing conditions (salt concentration: 0.2 ~ 0.5 X SSC, temperature: room temperature ~ 42˚C). As a result, we obatined a clear band of ~ 5kb digested by EcoRI when we perform hybridization overnight at 55˚C and wash twice with 2 X SSC at room temperature for 30 min, then wash twice more with 0.5 X SSC at 42˚C for 30 min. Thus it suggested the presence of OCH1-homologous gene on Pichia choromosome by the hybridization technique involving the mild hybridization temperature and washing condition.

[0051] (4) λgt10 library preparation
Following the result in (3), we attempted cloning of EcoRI fragment (~5kb) from Pichia chromosomal DNA. First after digesting ~150 μg chromosomal DNA derived from P. pastoris GTS115 strain with 200 enzyme units of EcoRI, we separated DNA ranging around 4.5 ~ 6 kb on the 0.8% agarose gel. Then we ligated the part of these DNA with 1 μg of λgt10arm ("lambda gt10 vector digested with EcoRI and dephosphorylated," Stratagene Inc.) and packaged using Gigapack II Gold Packaging Extract (Stratagene Inc.). We obtained a number of plaques enough for screening.

[0052] (5) Plaque hybridization
After titrating the recombinant λ phage library prepared in (4) to 200 to 300 plaques per 80mm-plate, we transferred it to the nylon membrane filter (Hybond-N, Amersham Inc.). By preparing 10 filters (total screening number: ~ 3000 plaques), we performed hybridization with the OCH1 gene fragment as a probe. As a result, we detected 14 clear positive plaques.

[0053] (6) λDNA isolation
After randomly selecting 10 plaques among positive ones detected in (5), we performed single plaque isolation and purified λDNA using Sephaglas™ PhagePrep Kit (Pharmacia Inc.). Each λDNA was digested with various restriction enzymes (EcoRI, BglII, HindIII, XhoI) and it was found by agarose gel electrophoresis that 8 clones out of 10 possess the identical inserted DNA (~5 kb).

[0054] (7) Subcloning
We picked up one clone and subcloned the EcoRI insert fragment into EcoRI site of pUC19 vector, resulting in pKM50 (See Fig. 4).

[0055] (8) Determination of DNA sequence and amino acid sequence inserted in pKM50
We determined the nucleotide sequence of Pichia-derived EcoRI fragment inserted in pKM50. By constructing the restriction map of cloned DNA fragment in pKM50 and performing Southern analysis with more restriction enzymes, it was shown that there is a homologous region with S. cerevisiae-derived OCH1 gene in the BglII fragment of ~2.5 kb. We determined the nucleotide sequence of this BglII fragment. To put it concretely, the fragment was cut with several restriction enzyme and subloned into pUC19, then sequenced with M13-40 primer and reverse primer (Pharmacia LKB Biotechnology) by using the DNA sequencer (A.L.F. DNA sequencer,

Pharmacia LKB Biotechnology).

[0056] We sequenced the OCH1-homologous DNA fragment [Bgl II – Sal I fragment (~3.0 kb)] and found the open reading frame (ORF, oblique box region in Fig. 4) consisting of 404 amino acids. We show the nucleotide sequence from Bgl II to Sal I site(2858bp) that includes ORF region in Sequence Table 1. The region possesses two potential sites of asparagine-binding glycosylation (Asn-Xaa-Ser/Thr).

[0057] We compared the Pichia-derived ORF amino acid sequence described above with that of S. cerevisiae OCH1 protein. We found that it has an identity of about 40% (see Fig. 5). At the nucleotide level that encodes the amino acid, the identity was about 55%. Among the putative asparagine-binding glycosylation sites (boxed in Fig. 5), one site was homologous each other (Asn^{199} in Pichia-derived protein and Asn^{203} in S. cerevisiae OCH1 protein). The expected molecular weight of Pichia-derived protein was 46 kDa while that of S. cerevisiae OCH1 is 55 kDa.

[0058] Next we compared the hydrophobicity of both proteins. As is shown in Fig. 6, both showed very similar pattern. Therefore the DNA we obtained was suggested to be OCH1 gene derived from the Pichia strain. Also we found that S. cerevisiae OCH1 protein possesses hydrophobic region of putative transmembrane domain around the N-terminus and that Pichia-derived protein possesses even longer hydrophobic region.

[0059] Working example 2: Preparation of glycosylation gene-disrupted strains
(5) Genomic southern analysis of Pichia-derived glycosylation DNA
For the purpose of making strains in which glycosylation DNA we cloned in working example 1 is deleted, we first performed genomic southern hybridization to confirm the interested DNA is a single gene on chromosome. Chromosomal DNA from host cell P. pastoris GTS115 strain was digested with either Bgl II, EcoRI, Sph I, or Xba I, electorophoresed on agarose gel, and blotted onto the nylon membrane. Next we prepared a Pichia-derived DNA fragment corresponding to nt. 1488 through 2385 described in Sequence Table 1 and conducted hybridization with this fragment as a probe. The result is shown in Fig. 7. It indicates that the probe hybridize only a single band of DNA digested with any of the restriction enzymes, thereby confirming that the Pichia-derived glycosylation DNA is a single gene.

[0060] (2) Preparation of mutants that have HIS4 as a selective marker
pKM50 (Fig. 4), which contains Pichia-derived glycosylation DNA and its surrounding nucleotides, was blunt-end digested with Asu II and Bal I, and inserted with the expression unit comprising of HIS4 gene and soluble high-affinity IgE receptor alpha chain gene (sFcɛRIα), producing plasmid pKM74 (See Fig. 8). The sFcɛRIα expression unit possess a signal sequence from S. cerevisiae SUC2 gene in N-terminus of sFcɛRIα sequence [Nucleic Acids Research, volume 16, Number 8, 3584 (1988)], and contains P. pastoris AOX2 gene promotor and P. pastoris AOX1 gene terminator; it is capable of secreting extracellular region (172 a.a.) of human-derived high-affinity IgE receptor in P. pastoris.

[0061] After digesting pKM74 with Sph I and Pst I and transforming P. pastoris GTS115 strain (his4) [NRRL Y-15851], we obtained 45 transformants (HIS4). Picking up a couple of them, we performed the following analyses.

[0062] (3) Analysis of GTS115/pKM74 transformants

It was reported that the S. cerevisiae-derived OCH1 disruptant lost resistance to high temperature and it cannot be grown at 37˚C [Nakayama K et al. EMBO J. 11, 2511 (1992)]. We therefore examined the temperature sensitivity of the transformants we obtained in (2). We incubated 45 transformants on YPD plates at 20˚C, 30˚C, or 37˚C, and observed that 10 lines was not able to grow at 37˚C. On the other hand, we randomly picked 10 transformants and performed genomic southern hybridization, finding that 2 lines (KM74-2 and KM74-5) lack the glycosylation DNA. Both of the two lines were unable to grow at 37˚C. Thus the results of temperature sensitivity and southern hybridization were consistent each other.

[0063] Furthermore, we performed more detailed genomic southern analysis of one transformant (KM74-2) [Fig. 9]. KM45 strain shown in Fig. 9 is a transformant where the expression unit consisting of pKM74-derived HIS4 and sFcɛRIα gene is inserted into his4 locus in P. pastoris GTS115 his4 strain; it is capable of secreting sFcɛRIα protein. We performed genomic southern analysis on GTS115 strain, OCH1 wild type strain KM45, and OCH1-deletion strain KM74-2 by using as probes the upstream region (Probe 1 in Fig. 9: Bgl II – Asu II fragment in pKM50 shown in Fig. 4, 1256bp, nt. sequence 2 – 1258) and the internal region (Probe 2 in Fig. 9: Hind III – EcoT14I fragment in pKM 50 in Fig. 4) of glycosylation DNA. We confirmed that glycosylation DNA from KM74-2 is disrupted by inserted pKM74 DNA fragment.

[0064] (4) Analysis of sFcɛRIα protein which glycosylation DNA disruptants produce
In order to look at the glycosylation by Pichia-derived strain in which glycosylation DNA is disrupted, we first cultured the disruptant KM74-2 and KM74-5, and the wild-type KM45 strain in 3X YP + 2% methanol M medium (3% yeast extract, 6% bactopeptone, 2% methanol) at 25˚C for 4 days. sFcɛRIα protein was then purified from culture supernatant by IgE affinity column. We analyzed by SDS polyacrylamide gel electrophoresis the untreated sFcɛRIα protein and PNGaseF (Genzyme Inc.) -treated protein in which asparagine-binding sugar chain is removed (Fig. 12). The result showed that while high molecular weight of sFcɛRIα was observed in wild-type KM45 (Fig. 12, lane 1), its molecular weight was lower in disruptants KM74-2 and KM74-5 (Fig. 12, lane 2 & 3). When the sugar chain is removed by PNGaseF, proteins from all the strains showed the same molecular weight, demonstrating that the difference in molecular weight comes from glycosylation. These result indicate that glycosylation is blocked in P. pastrosis-derived disruptants.

[0065]
[Sequence Table]
Sequence number: 1
Length of sequence: 2858
Sequence classification: nucleic acid
Number of chains: 2
Topology: straight
Sequence type: genomic DNA
Organism of origin: P. pastoris
Strain: GTS115
Sign showing character: CDS
Position: 1027 – 2238
Method of characterization: S, P

(Sequence)

[Concise figure legend]

[Fig. 1] Diagram showing functional model of Asn-binding sugar chain in erythropoietin. Man; mannose, GlcNAc; N-acetylglucosamine, Fuc; fucose

[Fig. 2] Diagram showing the structure of glycosylation in glycoprotein in S. cerevisiae. "Ma" in N-linked sugar chain indicates mannose chain synthesized in ER. M; mannose, 2; α-1,2 binding, 6; α-1,6 binding, 4; α-1,4 binding

[Fig. 3] Plasmid map showing pKM049 in which S. cerevisiae-derived OCH1 gene is subcloned.

[Fig. 4] Restriction map of DNA fragment inserted into pKM50 (P. pastoris-derived chromosomal DNA that possesses homology with S. cerevisiae-derived OCH1 gene). The boxed region with oblique lines indicates putative translating region of OCH1 protein.

[Fig. 5] Figure showing amino acid sequence homology between the protein encoded by S. cerevisiae OCH1 gene and putative glycosylation protein encoded by P. pastoris DNA. Box shows asparagine chain addition site.

[Fig. 6] Comparison of hydrophobicity profiles between S. cerevisiae-derived OCH1 protein (A) and P. pastoris-derived glycosylation protein (B).

[Fig. 7] Picture of agarose gel electrophoresis showing the result of genomic southern hybridization using Pichia-derived glycosylation DNA as a probe.

[Fig. 8] Restriction map illustrating the plasmid (pKM74) in which P. pastoris-derived glycosylation DNA is disrupted.

[Fig. 9] Structure of the glycosylation gene locus and its surroundings in chromosomal DNA from the disruptant KM74-2, and wild-type GTS115 and KM45. The underline shows the position of the probe used in the genomic southern hybridization analysis. E; EcoRI, Bg; Bgl II

[Fig. 10] Picture showing the result of genomic southern hybridization analysis for the disruptant KM74, and wild-type GTS115 and KM45 by using the probe 1 shown in Fig. 9.

[Fig. 11] Picture showing the result of genomic southern hybridization analysis for the disruptant KM74, and wild-type GTS115 and KM45 by using the probe 2 shown in Fig. 9.

[Fig. 12] SDS-PAGE analysis of sFcεRIα protein produced by P. pastoris-derived disruptants. 1: sFcεRIα (KM45), 2: sFcεRIα (KM74-2), 3: sFcεRIα (KM74-5), 4: PNGaseF-treated sFcεRIα (KM45), 5: PNGaseF-treated sFcεRIα (KM74-2), 6: PNGaseF-treated sFcεRIα (KM74-5)

(Figure 1 – 12)